

# PHASE RETRIEVAL FROM POWER SPECTRA OF MASKED SIGNALS

AFONSO S. BANDEIRA, YUTONG CHEN, AND DUSTIN G. MIXON

**ABSTRACT.** In diffraction imaging, one is tasked with reconstructing a signal from its power spectrum. To resolve the ambiguity in this inverse problem, one might invoke prior knowledge about the signal, but phase retrieval algorithms in this vein have found limited success. One alternative is to create redundancy in the measurement process by illuminating the signal multiple times, distorting the signal each time with a different mask. Despite several recent advances in phase retrieval, the community has yet to construct an ensemble of masks which uniquely determines all signals and admits an efficient reconstruction algorithm. In this paper, we leverage the recently proposed polarization method to construct such an ensemble. We also present numerical simulations to illustrate the stability of the polarization method in this setting. In comparison to a state-of-the-art phase retrieval algorithm known as PhaseLift, we find that polarization is much faster with comparable stability.

## 1. INTRODUCTION

In many applications, one wishes to reconstruct a signal from the magnitudes of its Fourier coefficients. This problem is known as *phase retrieval*, and it has been instrumental to many important scientific advances, including the Nobel Prize-winning work that leveraged X-ray diffraction to establish the double helix structure of DNA [22]. Of course, given only the magnitudes of a signal's Fourier coefficients, one does not have enough information to recover the signal—while the Fourier transform is injective, the point-wise absolute value is not. As such, one is inclined to use a priori knowledge of the signal, and hope it is then uniquely determined by the Fourier magnitudes. For example, to deduce the structure of DNA, Watson and Crick [22] applied certain chemical assumptions, along with a knowledge of van der Waals interactions between atoms.

In order to image more exotic molecules, such assumptions are difficult to apply, and so there has been quite a bit of work attempting to exploit more general assumptions (e.g., positivity or support constraints). To account for this sort of prior information, the most popular phase retrieval algorithms are modifications of Gerchberg and Saxton's original approach [13], which alternates between the time and frequency domains, iteratively correcting the current guess by imposing time-domain assumptions or scaling Fourier coefficients to match the measured data. Marchesini [17] surveys and compares the various modifications, but they all have a tendency to stall in local minima.

To address this issue, Candès, Eldar, Strohmer and Voroninski [7] proposed an alternative methodology whereby nonuniqueness is overcome not by prior information, but by additional illuminations. For each illumination, a different mask (or grating) is used to distort the appearance of the object in question; in mathematical parlance, each mask acts as a multiplication operator on the desired signal before the Fourier transform. Furthermore, Candès et al. constructed three masks such that the corresponding illuminations uniquely determine almost every signal up to a global phase factor; a fourth illumination is necessary to uniquely determine all signals [5]. However, the phase retrieval algorithms they propose

are not guaranteed to reconstruct the desired signal from these three illuminations, even if the signal is uniquely determined by the illuminations.

Let  $F^*: \ell(\mathbb{Z}_M) \rightarrow \ell(\mathbb{Z}_M)$  denote the discrete Fourier transform (DFT) defined by

$$(F^*x)(m) := \frac{1}{M} \sum_{m' \in \mathbb{Z}_M} x(m') e^{-2\pi i m m' / M} \quad \forall m \in \mathbb{Z}_M.$$

In this paper, we follow the lead of [7] to find masks  $\{D_k\}_{k=0}^{K-1}$  (i.e., multiplication operators on  $\ell(\mathbb{Z}_M) = \mathbb{C}^M$ ) such that every signal  $x \in \ell(\mathbb{Z}_M)$  can be efficiently reconstructed, up to a global phase factor, from measurements of the form

$$\{|F^* D_k^* x|^2\}_{k=0}^{K-1} = \{|(D_k F)^* x|^2\}_{k=0}^{K-1} = \{|\langle x, D_k f_m \rangle|^2\}_{k=0}^{K-1}, \quad m=0, \dots, M-1,$$

where  $f_m$  denotes the complex sinusoid  $\{\frac{1}{M} e^{2\pi i m m' / M}\}_{m' \in \mathbb{Z}_M}$ . Specifically, we will design the masks in such a way that allows for a new reconstruction technique to work, namely the polarization-based technique introduced in [1]. As established in [1], every signal  $x$  can be stably reconstructed from  $|\langle x, \varphi_i \rangle|^2$  if the measurement vectors  $\varphi_i$  are constructed using a particular random process. In this paper, we show how to fulfill the design criteria that measurement vectors be of the form  $D_k f_m$  while still allowing for polarization-based reconstruction. If anything, this demonstrates that polarization is flexible enough to provably accommodate certain measurement design requirements, which are bound to arise from various applications.

In the next section, we briefly review the main ideas behind the polarization-based technique, and we show how to apply them with Fourier masks. In particular, we reduce the problem of building masks to a problem in additive combinatorics: Find a small subset of  $\mathbb{Z}_M$  with small Fourier bias. In Section 3, we use concentration-of-measure arguments to solve this problem, culminating in the construction of  $\mathcal{O}(\log M)$  masks that uniquely determine every signal in  $\mathbb{C}^M$  up to a global phase factor. Next in Section 4, we compare the performance of polarization to a state-of-the-art phase retrieval algorithm known as PhaseLift [7, 9]. Here, we show that polarization provides a comparable amount of stability with Fourier masks, while being orders of magnitude faster. We conclude in Section 5 with some remarks.

## 2. HOW TO POLARIZE FOURIER MASKS

The main contribution of this paper is that the ideas in [1] can be leveraged to uniquely measure signals with masked Fourier transforms. In this section, we summarize these ideas and explain how we apply them. In [1], the signal  $x$  is measured with two ensembles of vectors  $\Phi_V, \Phi_E \subseteq \mathbb{C}^M$ . Here,  $\Phi_V$  has the property that every subcollection of size  $M$  spans  $\mathbb{C}^M$ ; such ensembles are called *full spark frames*, of which there are several known constructions [2, 18]. To determine  $\Phi_E$ , first define a graph  $G = (V, E)$  whose vertices  $i \in V$  index the measurement vectors  $\varphi_i \in \Phi_V$ . Then each edge  $(i, j) \in E$  corresponds to three members of  $\Phi_E$ , namely  $\{\varphi_i + \omega^r \varphi_j\}_{r=0}^2$ , where  $\omega = e^{2\pi i / 3}$ . These edges are useful because they determine the relative phase between the inner products  $\langle x, \varphi_i \rangle$  and  $\langle x, \varphi_j \rangle$ . Specifically, a version of the polarization identity gives that

$$\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle = \frac{1}{3} \sum_{r=0}^2 \omega^r |\langle x, \varphi_i \rangle + \omega^{-r} \langle x, \varphi_j \rangle|^2 = \frac{1}{3} \sum_{r=0}^2 \omega^r |\langle x, \varphi_i + \omega^r \varphi_j \rangle|^2, \quad (1)$$

and so the relative phase is calculated by normalizing this quantity (provided it's nonzero). On the other hand, if an inner product  $\langle x, \varphi_i \rangle$  is zero, then for every vertex  $j$  adjacent to  $i$  in  $G$ , the members of  $\Phi_E$  corresponding to  $(i, j)$  are useless, since the relative phases they are intended to capture are not well defined. Indeed,  $x$  being orthogonal to  $\varphi_i$  has the effect of deleting the vertex  $i$  (and its edges) from the graph  $G$ . However, since  $\Phi_V$  is full spark,  $x$  is orthogonal to at most  $M - 1$  vertex vectors, thereby limiting the damage to our graph. Furthermore, for any remaining connected component  $S \subseteq V$  of size  $\geq M$ , we can propagate the relative phases to determine  $\{\langle x, \varphi_i \rangle\}_{i \in S}$  up to a global phase factor, and since  $\{\varphi_i\}_{i \in S}$  necessarily span  $\mathbb{C}^M$ , we can determine  $x$  from  $\{\langle x, \varphi_i \rangle\}_{i \in S}$  by least-squares estimation. Interestingly, there will always be a connected component of size  $\geq M$  after the removal of any  $M - 1$  vertices, provided the graph  $G$  is regular with a sufficiently large *spectral gap*. To be clear, the spectral gap of a  $d$ -regular graph  $G$  is determined by the eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots$  of its adjacency matrix:

$$\text{spg}(G) = \frac{1}{d} \left( \lambda_1 - \max_{i \neq 1} |\lambda_i| \right),$$

and the existence of the requisite connected component is given by the following classical result in spectral graph theory (e.g., see Lemma 5.2 of [15]):

**Lemma 1.** *Consider a  $d$ -regular graph  $G$  of  $n$  vertices. For all  $\varepsilon \leq \text{spg}(G)/6$ , removing any  $\varepsilon dn$  edges from  $G$  results in a connected component of size  $\geq (1 - 2\varepsilon/\text{spg}(G))n$ .*

In summary, in order to apply the polarization trick in the setting of masked DFTs, it suffices to (i) construct a full spark ensemble  $\Phi_V$  with masked DFTs, and (ii) construct a graph  $G$  with a sufficiently large spectral gap and whose corresponding edge vectors  $\Phi_E$  can also be implemented with masked DFTs. We quickly address (i) with the following result:

**Lemma 2.** *For each  $k \in \{0, \dots, K - 1\}$ , pick some nonzero  $\alpha_k \in \mathbb{C}$  and define  $D_k := \text{diag}\{\alpha_k^m\}_{m=0}^{M-1}$ . Let  $f_m$  denote the complex sinusoid  $\{\frac{1}{M}e^{2\pi i m m'/M}\}_{m' \in \mathbb{Z}_M}$ . Then  $\{D_k f_m\}_{k=0, m=0}^{K-1, M-1}$  is full spark if and only if  $\alpha_k \alpha_{k'}^{-1}$  is not an  $m$ th root of unity for any pair of distinct  $k, k' \in \{0, \dots, K - 1\}$ .*

*Proof.* The  $M \times KM$  matrix whose columns are  $D_k f_m$  has Vandermonde form, as does each of its  $M \times M$  submatrices. These submatrices are all invertible precisely when their determinants are nonzero, that is, when  $\{\alpha_k e^{2\pi i m/M}\}_{k=0, m=0}^{K-1, M-1}$  are distinct (by the Vandermonde determinant formula). The lemma immediately follows.  $\square$

The condition above is satisfied with probability 1 if the  $\alpha_k$ 's are drawn uniformly at random from the complex unit circle. For a deterministic alternative, it suffices to take  $\alpha_k = e^{2\pi i k/KM}$ . Now that we have established how to construct masks  $D_k$  such that  $\Phi_V = \{D_k f_m\}_{k=0, m=0}^{K-1, M-1}$  is full spark, we turn to solving (ii). Before describing our construction of  $G$ , we first motivate it by illustrating how polarized combinations of our vertex vectors can be expressed using masked DFTs:

**Lemma 3.** *Take  $\omega = e^{2\pi i/3}$ , let  $E = \text{diag}\{e^{2\pi i m/M}\}_{m=0}^{M-1}$  denote the modulation operator, and consider  $\Phi_V = \{D_k f_m\}_{k=0, m=0}^{K-1, M-1}$ , as defined in Lemma 2. Then*

$$D_k f_m + \omega^r D_{k'} f_{m'} = (D_k + \omega^r E^{m'-m} D_{k'}) f_m$$

*for all  $k, k' \in \{0, \dots, K - 1\}$  and  $m, m' \in \mathbb{Z}_M$ .*

*Proof.* Consider the  $\ell$ th entry of  $D_{k'} f_{m'}$ :

$$(D_{k'} f_{m'}) (\ell) = \alpha_{k'}^\ell \cdot \frac{1}{M} e^{2\pi i m' \ell / M} = e^{2\pi i (m' - m) \ell / M} \cdot \alpha_{k'}^\ell \cdot \frac{1}{M} e^{2\pi i m \ell / M} = (E^{m' - m} D_{k'} f_m) (\ell).$$

Then  $D_k f_m + \omega^r D_{k'} f_{m'} = D_k f_m + \omega^r E^{m' - m} D_{k'} f_m = (D_k + \omega^r E^{m' - m} D_{k'}) f_m$ .  $\square$

When implementing the modulation trick of Lemma 3 using DFTs, we will have to fix  $m' - m$  (so as to apply a fixed mask) and let  $m$  vary (since we will mask an entire DFT). As such, we intend to use auxiliary masks of the form  $\{D_k + \omega^r E^a D_{k'}\}_{r=0}^2$ , thereby drawing an edge between every  $(k, m)$  and  $(k', m')$  such that  $m' - m = a \pmod{M}$ . This informs our decision of how to construct the graph  $G$ , since it enables a masked-DFT implementation:

**Definition 4.** *We construct  $G$  as follows: Pick  $A \subseteq \mathbb{Z}_M$  such that  $0 \notin A$  and  $A = -A$ . Then  $(k, m)$  and  $(k', m')$  are adjacent precisely when  $m' - m \in A$ .*

Note that for each  $a \in A$ , the corresponding edges between  $\{(k, m)\}_{m \in \mathbb{Z}_M}$  and  $\{(k', m')\}_{m' \in \mathbb{Z}_M}$  are implemented with the masks  $\{D_k + \omega^r E^a D_{k'}\}_{r=0}^2$ . As such,  $A$  is the set of modulations used to combine the  $K$  original masks  $D_k$  into  $3 \binom{K+1}{2} |A|$  auxiliary masks. Also note that  $G$  has no loops since  $0 \notin A$ ,  $G$  is not directed since  $A = -A$ , and furthermore every vertex has degree  $K|A|$ , so  $G$  is regular. It remains to show that  $G$  has a sufficiently large spectral gap. To simplify this analysis, we first relate the spectral gap to a fundamental concept in additive combinatorics called the *Fourier bias*. As in [20], take  $A \subseteq \mathbb{Z}_M$  and let  $\mathbf{1}_A$  denote the characteristic function of  $A$ . Then the Fourier bias of  $A$  is defined as follows:

$$\|A\|_u := \max_{m \neq 0} |(F^* \mathbf{1}_A)(m)|.$$

Fourier bias is used in additive combinatorics to measure pseudorandomness; here, the main idea is that correlation with any complex sinusoid indicates regularity, which is not typically exhibited by random sets. As indicated earlier, the spectral gap of  $G$  is intimately related to the Fourier bias of  $A$ :

**Lemma 5.** *Consider the graph  $G$  defined in Definition 4. The spectral gap of  $G$  is*

$$\text{spg}(G) = 1 - \frac{M}{|A|} \|A\|_u,$$

where  $\|\cdot\|_u$  denotes Fourier bias.

*Proof.* To determine the spectral gap, we first determine the adjacency matrix  $W$  of  $G$ . For any pair  $k, k' \in \{0, \dots, K-1\}$ , the adjacency rule for  $(k, m)$  and  $(k', m')$  is whether  $m' - m \in A$ . As such, the  $(k, k')$ th block of  $W$  is circulant, with each row being a translation of  $\mathbf{1}_A$ . This gives the expression  $W = J \otimes \text{circ}(\mathbf{1}_A)$ , where  $J$  is the  $K \times K$  all-ones matrix and  $\otimes$  denotes the Kronecker product. One useful property of the Kronecker product is that its eigenvalues are products of eigenvalues:

$$\lambda_{(i,j)}(W) = \lambda_i(J) \lambda_j(\text{circ}(\mathbf{1}_A)).$$

Here, we note that the only nontrivial eigenvalue of  $J$  is  $K$ , and the eigenvalues of  $\text{circ}(\mathbf{1}_A)$  are the entries of the Fourier transform  $M F^* \mathbf{1}_A$ . As such, the nontrivial eigenvalues of  $W$  are given by

$$\lambda_{(1,m)}(W) = \lambda_1(J) \lambda_m(\text{circ}(\mathbf{1}_A)) = KM(F^* \mathbf{1}_A)(m) = K \sum_{m' \in \mathbb{Z}_M} \mathbf{1}_A(m') e^{-2\pi i m m' / M}.$$

By the triangle inequality, we then have

$$|\lambda_{(1,m)}(W)| = K \left| \sum_{m' \in \mathbb{Z}_M} \mathbf{1}_A(m') e^{-2\pi i m m' / M} \right| \leq K \sum_{m' \in \mathbb{Z}_M} |\mathbf{1}_A(m')| = K|A|,$$

with equality when  $m = 0$ . This means the largest eigenvalue of  $W$  is  $\lambda_1 = K|A|$ , which corresponds to the all-ones eigenvector, and so the spectral gap of  $G$  is

$$\text{spg}(G) = \frac{1}{d} \left( \lambda_1 - \max_{i \neq 1} |\lambda_i| \right) = \frac{1}{K|A|} \left( K|A| - \max_{m \neq 0} |KM(F^* \mathbf{1}_A)(m)| \right) = 1 - \frac{M}{|A|} \|A\|_u,$$

as claimed.  $\square$

At this point, we have reduced the problem of finding injective Fourier masks to finding a small set  $A \subseteq \mathbb{Z}_M$  of sufficiently small Fourier bias (i.e., a problem in additive combinatorics). In the next section, we use the following random process to construct this set: Draw  $B \subseteq \mathbb{Z}_M$  at random and take  $A := B \cup (-B) \setminus \{0\}$ . With the appropriate choice of distribution, we construct  $A$  of size  $\Theta(\log M)$  and Fourier bias  $\mathcal{O}(\frac{\log M}{M})$ , thereby producing a graph  $G$  with a spectral gap that is bounded away from zero (see Theorem 9). We also show that the spectral gap is bounded away from zero only if  $A$  has size  $\Omega(\log M)$ , thereby demonstrating the optimality of our analysis (see Theorem 11). Overall, letting  $K$  be a sufficiently large constant, the techniques of this paper use a total of  $K + 3 \binom{K+1}{2} |A| = \Theta(\log M)$  Fourier masks to uniquely determine any signal (see Theorem 10).

### 3. SMALL SETS WITH SMALL FOURIER BIAS

In the previous section, we reduced the main problem of this paper to one in additive combinatorics: Find a small set  $A \subseteq \mathbb{Z}_M$  with small Fourier bias  $\|A\|_u = \max_{m \neq 0} |(F^* \mathbf{1}_A)(m)|$ . In our case, we want  $A$  to satisfy two additional properties:  $0 \notin A$  and  $A = -A$ . To account for these, we will approach this problem by first drawing  $B \subseteq \mathbb{Z}_M$  at random, and then taking  $A := B \cup (-B) \setminus \{0\}$ . We start with the following lemma:

**Lemma 6.** *Suppose the entries of  $\mathbf{1}_B$  are independent, identical Bernoulli random variables with mean  $\frac{c \log M}{M}$ . Then the following simultaneously hold with high probability:*

- (i)  $B \cap (-B) = \emptyset$
- (ii)  $0 \notin B$
- (iii)  $\frac{1}{2} c \log M \leq |B| \leq \frac{3}{2} c \log M$

*Proof.* First, a union bound gives

$$\begin{aligned} \Pr(B \cap (-B) \neq \emptyset) &= \Pr(\exists m \in \mathbb{Z}_M \text{ s.t. } m, -m \in B) \\ &\leq M \Pr(m \in B) \Pr(-m \in B) = \frac{c^2 \log^2 M}{M}. \end{aligned}$$

Next, we have  $\Pr(0 \in B) = \frac{c \log M}{M}$ . For (iii), we apply a multiplicative form of the Chernoff bound (see (6) and (7) in [14]): Let  $X$  be a sum of independent 0-1 random variables. Then

$$\Pr(|X - \mathbb{E}[X]| \geq \delta \mathbb{E}[X]) \leq 2e^{-\delta^2 \mathbb{E}[X]/3}$$

for any choice  $0 < \delta < 1$ . Here, we let  $X := \sum_{m \in \mathbb{Z}_M} \mathbf{1}_B(m) = |B|$  and take  $\delta = 1/2$ . Then

$$\Pr\left(|B| - c \log M \geq \frac{1}{2} c \log M\right) \leq 2e^{-(c \log M)/12} = 2M^{-c/12}.$$

The result then follows from a union bound.  $\square$

Note that in the event of Lemma 6, we have  $\mathbf{1}_A = \mathbf{1}_B + \mathbf{1}_{-B}$ , and so

$$\begin{aligned} \|A\|_u &= \max_{m \neq 0} |(F^* \mathbf{1}_A)(m)| \leq \max_{m \neq 0} \left( |(F^* \mathbf{1}_B)(m)| + |(F^* \mathbf{1}_{-B})(m)| \right) \\ &\leq \max_{m \neq 0} |(F^* \mathbf{1}_B)(m)| + \max_{m \neq 0} |(F^* \mathbf{1}_{-B})(m)| = 2\|B\|_u, \end{aligned} \quad (2)$$

where the last step applied a complex conjugate to  $(F^* \mathbf{1}_{-B})(m)$ . As such, it suffices to show that random sets  $B$  have small Fourier bias:

**Lemma 7.** *Pick  $c \geq 4$  and suppose the entries of  $\mathbf{1}_B$  are independent, identical Bernoulli random variables with mean  $\frac{c \log M}{M}$ . Then  $\|B\|_u \leq 3\sqrt{c} \cdot \frac{\log M}{M}$  with high probability.*

To prove this lemma, we will apply the following version of the Chernoff bound:

**Lemma 8.** *Assume that  $\{Z_i\}_{i=1}^n$  are jointly independent complex random variables where  $|Z_i - \mathbb{E}[Z_i]| \leq 1$  for all  $i$ . Take  $Z := \sum_{i=1}^n Z_i$  and denote  $\sigma := \sqrt{\text{Var}(Z)}$ . Then*

$$\Pr \left( |Z - \mathbb{E}[Z]| \geq t\sigma \right) \leq 4 \max \left\{ e^{-t^2/8}, e^{-t\sigma/2\sqrt{2}} \right\}$$

for any  $t > 0$ .

*Proof.* Let  $X$  and  $Y$  denote the real and imaginary parts of  $Z$ . Then a union bound gives

$$\begin{aligned} \Pr \left( |Z - \mathbb{E}[Z]| \geq t\sigma \right) &= \Pr \left( |X - \mathbb{E}[X]|^2 + |Y - \mathbb{E}[Y]|^2 \geq t^2 \sigma^2 \right) \\ &\leq \Pr \left( |X - \mathbb{E}[X]| \geq \frac{t\sigma}{\sqrt{2}} \right) + \Pr \left( |Y - \mathbb{E}[Y]| \geq \frac{t\sigma}{\sqrt{2}} \right). \end{aligned}$$

Denote  $\sigma_X := \sqrt{\text{Var}(X)}$  and  $\sigma_Y := \sqrt{\text{Var}(Y)}$ . Since  $\text{Var}(X) + \text{Var}(Y) = \text{Var}(Z)$ , we then have  $\sigma_X \leq \sigma$  and  $\sigma_Y \leq \sigma$ , and so

$$\Pr \left( |Z - \mathbb{E}[Z]| \geq t\sigma \right) \leq \Pr \left( |X - \mathbb{E}[X]| \geq \frac{t\sigma_X}{\sqrt{2}} \right) + \Pr \left( |Y - \mathbb{E}[Y]| \geq \frac{t\sigma_Y}{\sqrt{2}} \right).$$

The result then follows from applying Theorem 1.8 of [20] to both terms of the right-hand side.  $\square$

*Proof of Lemma 7.* Fix  $m \in \mathbb{Z}_M$ ,  $m \neq 0$ , and take  $Z_{m'} := \frac{1}{2} \mathbf{1}_B(m') e^{-2\pi i m m' / M}$  for each  $m' \in \mathbb{Z}_M$ . Then  $\{Z_{m'}\}_{m' \in \mathbb{Z}_M}$  are jointly independent with

$$|Z_{m'} - \mathbb{E}[Z_{m'}]| \leq |Z_{m'}| + |\mathbb{E}[Z_{m'}]| \leq 1.$$

As such,  $\{Z_{m'}\}_{m' \in \mathbb{Z}_M}$  satisfy the conditions of Lemma 8, and so we now seek the expected value and variance of

$$Z := \sum_{m' \in \mathbb{Z}_M} Z_{m'} = \frac{1}{2} \sum_{m' \in \mathbb{Z}_M} \mathbf{1}_B(m') e^{-2\pi i m m' / M}.$$

Take  $p := \frac{c \log M}{M}$  to be the probability that  $\mathbf{1}_B(m')$  is 1. Then for each  $S \subseteq \mathbb{Z}_M$ , the probability that  $B = S$  is  $p^{|S|}(1-p)^{M-|S|}$ . This leads to the following calculation:

$$\begin{aligned} \mathbb{E}[Z] &= \sum_{S \subseteq \mathbb{Z}_M} \left( \frac{1}{2} \sum_{m' \in \mathbb{Z}_M} \mathbf{1}_S(m') e^{-2\pi i m m' / M} \right) p^{|S|} (1-p)^{M-|S|} \\ &= \frac{1}{2} \sum_{\substack{S \subseteq \mathbb{Z}_M \\ S \neq \emptyset}} \sum_{m' \in S} e^{-2\pi i m m' / M} p^{|S|} (1-p)^{M-|S|} \\ &= \frac{1}{2} \sum_{m' \in \mathbb{Z}_M} e^{-2\pi i m m' / M} \sum_{S \supseteq \{m'\}} p^{|S|} (1-p)^{M-|S|}. \end{aligned}$$

From here, we apply the binomial theorem to get

$$\sum_{S \supseteq \{m'\}} p^{|S|} (1-p)^{M-|S|} = p \sum_{T \subseteq \mathbb{Z}_M \setminus \{m'\}} p^{|T|} (1-p)^{(M-1)-|T|} = p,$$

and so  $\mathbb{E}[Z] = 0$  by the geometric sum formula (recall that  $m \neq 0$  by choice). Next, we compute  $\text{Var}(Z) = \mathbb{E}[|Z - \mathbb{E}[Z]|^2] = \mathbb{E}[|Z|^2] = \mathbb{E}[Z \bar{Z}]$ :

$$\begin{aligned} \text{Var}(Z) &= \sum_{S \subseteq \mathbb{Z}_M} \left( \frac{1}{2} \sum_{m' \in \mathbb{Z}_M} \mathbf{1}_S(m') e^{-2\pi i m m' / M} \right) \left( \frac{1}{2} \sum_{m'' \in \mathbb{Z}_M} \mathbf{1}_S(m'') e^{2\pi i m m'' / M} \right) p^{|S|} (1-p)^{M-|S|} \\ &= \frac{1}{4} \sum_{m' \in \mathbb{Z}_M} \sum_{m'' \in \mathbb{Z}_M} e^{2\pi i m (m'' - m')} \sum_{S \supseteq \{m', m''\}} p^{|S|} (1-p)^{M-|S|}. \end{aligned}$$

At this point, we break the sum over  $m'' \in \mathbb{Z}_M$  into cases in which  $m'' = m'$  and  $m'' \neq m'$ :

$$\begin{aligned} \text{Var}(Z) &= \frac{1}{4} \sum_{\substack{m' \in \mathbb{Z}_M \\ m'' = m'}} e^{2\pi i m (m'' - m')} \sum_{S \supseteq \{m', m''\}} p^{|S|} (1-p)^{M-|S|} \\ &\quad + \frac{1}{4} \sum_{m' \in \mathbb{Z}_M} \sum_{\substack{m'' \in \mathbb{Z}_M \\ m'' \neq m'}} e^{2\pi i m (m'' - m')} \sum_{S \supseteq \{m', m''\}} p^{|S|} (1-p)^{M-|S|} \\ &= \frac{1}{4} M p + \frac{1}{4} \sum_{m' \in \mathbb{Z}_M} \sum_{\substack{m'' \in \mathbb{Z}_M \\ m'' \neq m'}} e^{2\pi i m (m'' - m')} \sum_{S \supseteq \{m', m''\}} p^{|S|} (1-p)^{M-|S|}. \end{aligned}$$

Here, another application of the binomial theorem gives

$$\sum_{S \supseteq \{m', m''\}} p^{|S|} (1-p)^{M-|S|} = p^2 \sum_{T \subseteq \mathbb{Z}_M \setminus \{m', m''\}} p^{|T|} (1-p)^{(M-2)-|T|} = p^2,$$

and so by the geometric sum formula, we have

$$\begin{aligned} \text{Var}(Z) &= \frac{1}{4} M p + \frac{1}{4} p^2 \sum_{m' \in \mathbb{Z}_M} \sum_{\substack{m'' \in \mathbb{Z}_M \\ m'' \neq m'}} e^{2\pi i m (m'' - m')} \\ &= \frac{1}{4} M p + \frac{1}{4} p^2 \sum_{m' \in \mathbb{Z}_M} (-1) = \frac{1}{4} M p - \frac{1}{4} M p^2 = \frac{1}{4} M p (1-p). \end{aligned}$$

Having calculated the expected value and variance of  $Z$ , we are now ready to use Lemma 8 (and the fact that  $Z = \frac{M}{2}(F^*\mathbf{1}_B)(m)$ ). A union bound gives

$$\begin{aligned} \Pr\left(\|B\|_u \geq \frac{2t\sigma}{M}\right) &\leq (M-1) \Pr\left(|(F^*\mathbf{1}_B)(m)| \geq \frac{2t\sigma}{M}\right) \\ &= (M-1) \Pr\left(|Z - \mathbb{E}[Z]| \geq t\sigma\right) \leq 4(M-1) \max\left\{e^{-t^2/8}, e^{-t\sigma/2\sqrt{2}}\right\}. \end{aligned}$$

We select  $t = (\sqrt{2c(1+\varepsilon)} \log M)/\sigma$  and simplify the exponents in our probability bound:

$$\frac{t^2}{8} = \frac{c(1+\varepsilon) \log^2 M}{Mp(1-p)} \geq \frac{c(1+\varepsilon) \log^2 M}{Mp} = (1+\varepsilon) \log M, \quad \frac{t\sigma}{2\sqrt{2}} = \frac{\sqrt{c(1+\varepsilon)}}{2} \log M.$$

With this choice of  $t$ , we continue:

$$\begin{aligned} \Pr\left(\|B\|_u \geq \sqrt{8c(1+\varepsilon)} \cdot \frac{\log M}{M}\right) &\leq 4(M-1) \max\left\{e^{-(1+\varepsilon) \log M}, e^{-(\sqrt{c(1+\varepsilon)}/2) \log M}\right\} \\ &\leq 4 \max\left\{M^{-\varepsilon}, M^{1-\sqrt{c(1+\varepsilon)}/2}\right\}. \end{aligned}$$

Since  $c \geq 4$ , we therefore have that  $\|B\|_u \leq 3\sqrt{c} \cdot \frac{\log M}{M}$  with high probability.  $\square$

We can now combine Lemmas 6 and 7 to produce a graph  $G$  of the form in Definition 4 with a large spectral gap:

**Theorem 9.** *Pick  $c \geq 4$  and suppose the entries of  $\mathbf{1}_B$  are independent, identical Bernoulli random variables with mean  $\frac{c \log M}{M}$ . Take  $A := B \cup (-B) \setminus \{0\}$  and define  $G$  according to Definition 4. Then*

$$\text{spg}(G) \geq 1 - \frac{6}{\sqrt{c}}$$

*with high probability.*

*Proof.* We will assume that the events of Lemmas 6 and 7 hold simultaneously, as they will with high probability. Then  $|A| = 2|B|$ , and furthermore by (2),  $\|A\|_u \leq 2\|B\|_u$ . Starting with Lemma 5, we then have

$$\text{spg}(G) = 1 - \frac{M}{|A|} \|A\|_u \geq 1 - \frac{M}{|B|} \|B\|_u \geq 1 - \frac{M}{\frac{1}{2}c \log M} \cdot 3\sqrt{c} \cdot \frac{\log M}{M} = 1 - \frac{6}{\sqrt{c}},$$

where the second inequality applies Lemmas 6(iii) and 7.  $\square$

We now apply Lemma 1 to show that  $\mathcal{O}(\log M)$  Fourier masks suffice for injectivity:

**Theorem 10.** *Take  $K = 12$ ,  $c = 144$ , suppose the entries of  $\mathbf{1}_B$  are independent, identical Bernoulli random variables with mean  $\frac{c \log M}{M}$ , and take  $A := B \cup (-B) \setminus \{0\}$ . Then with high probability, the  $\leq 2 \cdot 10^5 \cdot \log M$  masks*

$$\{D_k\}_{k=0}^{K-1} \cup \{D_k + \omega^r E^a D_{k'}\}_{k=0, k'=0, r=0, a \in A}^{K-1, 2}$$

*as described in Lemmas 2 and 3, lend injective intensity measurements.*

*Proof.* Note that  $G$  has  $n = KM$  vertices and is  $d$ -regular with  $d = K|A|$ . We need to be robust to the removal of any  $M-1$  vertices, which in turn removes  $d(M-1)$  edges, and so



it suffices to be robust to the removal of any  $dM$  edges. As such, we observe Lemma 1 and take  $\varepsilon$  to satisfy

$$\frac{1}{K} = \frac{dM}{dn} = \varepsilon \leq \frac{\text{spg}(G)}{6}. \quad (3)$$

We also want a connected component of size  $M$  after the removal of these edges, and so by Lemma 1, it suffices to have

$$M \leq \left(1 - \frac{2\varepsilon}{\text{spg}(G)}\right)n = \left(1 - \frac{2}{K \text{spg}(G)}\right)KM. \quad (4)$$

Rearranging (3) and (4) produces the following specification on the spectral gap of  $G$  for sufficient connectivity:

$$\text{spg}(G) \geq \max \left\{ \frac{6}{K}, \frac{2}{K-1} \right\} = \frac{6}{K},$$

where the equality is valid provided  $K \geq 2$ . To meet this specification, based on Theorem 9, it suffices to have  $1 - 6/\sqrt{c} \geq 6/K$ , which requires  $K > 6$  and is equivalent to having  $c \geq (1/6 - 1/K)^{-2}$ . Using Lemma 6, the total number of masks is

$$K + 3 \binom{K+1}{2} |A| = K + 6 \binom{K+1}{2} |B| \leq K + 9 \binom{K+1}{2} c \log M.$$

Setting  $c = (1/6 - 1/K)^{-2}$ , then  $K = 12$  minimizes the coefficient of  $\log M$ .  $\square$

At this point, we note that  $2 \cdot 10^5 \cdot \log M$  is a rather large number of Fourier masks. Certainly, the  $10^5$  might be an artifact of our analysis—perhaps it could be decreased by leveraging better approximations. However, as the next result shows, the log factor is necessary for the Fourier masks to lend polarization-based recovery:

**Theorem 11.** *Take  $G$  as defined in Definition 4. Then  $\text{spg}(G) > \varepsilon$  only if*

$$|A| \geq \frac{\log M}{2 + \log(1/\varepsilon)}.$$

*Proof.* Define  $V$  to be the  $|A| \times M$  matrix built from taking the rows of  $F$  indexed by  $A$  and then scaling the columns to have unit norm. Then the inner product between any two columns of  $V$  is given by

$$\begin{aligned} \langle v_m, v_{m'} \rangle &= \sum_{a \in A} \left( \frac{1}{\sqrt{|A|}} e^{2\pi i m a / M} \right) \left( \frac{1}{\sqrt{|A|}} e^{-2\pi i m' a / M} \right) \\ &= \frac{M}{|A|} \cdot \frac{1}{M} \sum_{m'' \in \mathbb{Z}_M} \mathbf{1}_A(m'') e^{-2\pi i (m' - m) m'' / M} = \frac{M}{|A|} (F^* \mathbf{1}_A)(m' - m). \end{aligned}$$

As such, the worst-case coherence between columns of  $V$  can be expressed in terms of the Fourier bias of  $A$  (and the spectral gap of  $G$  by Lemma 5):

$$\max_{\substack{m, m' \in \mathbb{Z}_M \\ m \neq m'}} |\langle v_m, v_{m'} \rangle| = \frac{M}{|A|} \|A\|_u = 1 - \text{spg}(G) \leq 1 - \varepsilon. \quad (5)$$

As you might expect, unit vectors can only be incoherent if there are enough dimensions relative to the number of vectors. To establish this explicitly, view the  $v_m$ 's as vectors in  $\mathbb{R}^{2|A|}$  by stacking the real and imaginary parts of each entry. Letting  $\delta$  denote the shortest distance between distinct  $v_m$ 's (this is the same in  $\mathbb{C}^{|A|}$  and  $\mathbb{R}^{2|A|}$ ), consider the open balls

of radius  $\delta/2$  centered at each  $v_m$ . The disjoint union of these  $M$  balls is contained in the ball of radius  $1 + \delta/2$  centered at the origin, and so a volume comparison gives

$$M \cdot C\left(\frac{\delta}{2}\right)^{2|A|} = \text{Vol} \left( \bigsqcup_{m \in \mathbb{Z}_M} B\left(v_m, \frac{\delta}{2}\right) \right) \leq \text{Vol} \left( B\left(0, 1 + \frac{\delta}{2}\right) \right) = C\left(1 + \frac{\delta}{2}\right)^{2|A|}.$$

Rearranging this inequality then yields

$$|A| \geq \frac{\log M}{2 \log(2/\delta + 1)}. \quad (6)$$

To analyze  $\delta$ , note that for any  $m, m' \in \mathbb{Z}_M$ , we have

$$\|v_m - v_{m'}\|^2 = \|v_m\|^2 - 2 \text{Re}\langle v_m, v_{m'} \rangle + \|v_{m'}\|^2 \geq 2 - 2|\langle v_m, v_{m'} \rangle| \geq 2\varepsilon,$$

where the last step is by (5). Thus  $\delta^2 \geq 2\varepsilon$ , with which we continue (6):

$$|A| \geq \frac{\log M}{2 \log(\sqrt{2/\varepsilon} + 1)} \geq \frac{\log M}{2 + \log(1/\varepsilon)}.$$

Here, the last inequality can be verified using the fact that  $\varepsilon \leq 1$ .  $\square$

#### 4. NUMERICAL SIMULATIONS

In this section, we compare the polarization method of this paper to a state-of-the-art phase retrieval algorithm known as *PhaseLift* [7, 9]. The main idea behind *PhaseLift* is that the intensity measurement  $x \mapsto |\langle x, \varphi \rangle|^2$  can be viewed as a linear measurement if the signal  $x$  is “lifted” to the outer product  $xx^*$  in the real vector space of self-adjoint  $M \times M$  matrices. Indeed, the intensity measurement is a Hilbert-Schmidt inner product in this vector space:

$$|\langle x, \varphi \rangle|^2 = \varphi^* x x^* \varphi = \text{Tr}[\varphi^* x x^* \varphi] = \text{Tr}[\varphi \varphi^* x x^*] = \langle x x^*, \varphi \varphi^* \rangle_{\text{HS}}.$$

However, this larger vector space has  $M^2$  dimensions, and so  $M^2$  inner products are necessary to identify any member of the space—that is, unless more information is available. In this case, we know that the desired self-adjoint matrix  $xx^*$  is positive semidefinite with rank 1, and so one could seek to minimize rank over the positive semidefinite matrices with the given intensity measurements. Rank minimization tends to be a difficult program to solve, so one is inclined to relax it:

$$(\text{PhaseLift}) \quad \text{minimize} \quad \text{Tr}(X) \quad \text{s.t.} \quad \mathcal{A}(X) = b, \quad X \succeq 0.$$

To date, *PhaseLift* is known to stably reconstruct all signals in  $\mathbb{C}^M$  with only  $\mathcal{O}(M)$  complex Gaussian measurements [8]. However, very little is known about its performance when the measurement vectors do not come from a unitarily invariant distribution (e.g., in the Fourier masks setting). To be fair, the present paper also does not prove stable reconstruction from Fourier masks; indeed, the previous section “merely” established injectivity with efficient recoverability. Therefore, for the sake of completeness, this section will provide numerical simulations that compare the stability of both phase retrieval algorithms.

First, we describe how we implement phase retrieval with polarization. Starting with an ensemble of vertex measurement vectors  $\Phi_V$  and edge measurement vectors  $\Phi_E$ , then for each edge  $(i, j) \in E$ , we apply (1) to get

$$w_{ij} = \frac{1}{3} \sum_{r=0}^2 \omega^r (|\langle x, \varphi_i + \omega^r \varphi_j \rangle|^2 + \varepsilon_{ijr}),$$

where  $\varepsilon_{ijr}$  denotes the noise added to the  $(i, j, r)$ th edge measurement. Later, we will normalize  $w_{ij}$  in order to estimate the relative phase between the measurements at vertices  $i$  and  $j$ , but if  $w_{ij}$  is small, this normalization will be particularly susceptible to noise. As such, we first remove vertices which are adjacent to small edge weights so as to promote reliability in the edges.

---

**Algorithm 1:** Pruning for reliability
 

---

**Input:** Weighted graph  $G = (V, E, W)$ , pruning parameter  $\alpha$

**Output:** Subgraph  $G'$  with a larger smallest edge weight

Initialize  $G' \leftarrow G$

**for**  $i = 1$  **to**  $\lfloor (1 - \alpha)|V| \rfloor$  **do**

    Find edge  $(i, j) \in E$  of minimal weight

$G' \leftarrow G' \setminus \{i, j\}$

**end**

---

Now that we have isolated a subgraph with reliable edges, we want to find a further subgraph which is sufficiently connected so that its vertices can democratically agree on the relative phases. To find this subgraph, we iteratively remove vertices implicated by *spectral clustering* [3, 4, 10] until the spectral gap is sufficiently large.

---

**Algorithm 2:** Pruning for connectivity
 

---

**Input:** Graph  $G' = (V', E')$ , pruning parameter  $\tau$

**Output:** Subgraph  $G''$  with  $\text{spg}(G'') \geq \tau$

Initialize  $G'' \leftarrow G'$

**while**  $\text{spg}(G'') < \tau$  **do**

    Take  $D$  to be the diagonal matrix of vertex degrees

    Compute the Laplacian  $L \leftarrow I - D^{-1/2}AD^{-1/2}$

    Compute the eigenvector  $u$  corresponding to the second eigenvalue of  $L$

**for**  $i = 1$  **to**  $|V''|$  **do**

        Let  $S_i$  denote the vertices corresponding to the  $i$  smallest entries of  $D^{-1/2}u$

        Let  $E(S_i, S_i^c)$  denote the number of edges between  $S_i$  and  $S_i^c$

$h_i \leftarrow E(S_i, S_i^c) / \min\{\sum_{v \in S_i} \deg(v), \sum_{v \in S_i^c} \deg(v)\}$

**end**

    Take  $S$  to be the  $S_i$  of minimal  $h_i$

$G'' \leftarrow G'' \setminus S$

**end**

---

At this point, our graph has reliable edges and is well connected. We now run *angular synchronization* [6, 19] to reach a consensus on the phases of the vertex measurements, up to a global phase factor.

Having estimated the phases of the vertex measurements corresponding to  $V''$ , we now multiply the square roots of the vertex measurements  $\{\sqrt{|\langle x, \varphi_i \rangle|^2 + \varepsilon_i}\}_{i \in V''}$  by these phases to estimate the inner products  $\{\langle x, \varphi_i \rangle\}_{i \in V''}$ , and then produce a least-squares estimate for the desired signal  $x$ . As established in [1], this implementation of the polarization method is stable to noise when the vertex measurement vectors are complex Gaussian. In this section, we run numerical simulations to illustrate stability in the Fourier masks setting.

**Algorithm 3:** Angular synchronization**Input:** Graph  $G'' = (V'', E'', W'')$ **Output:** Vector of phases corresponding to vertex measurementsLet  $A_1$  denote the weighted adjacency matrix with the weights normalizedLet  $D$  denote the diagonal matrix of vertex degreesCompute the connection Laplacian  $L_1 \leftarrow I - D^{-1/2} A_1 D^{-1/2}$ Compute the eigenvector  $u$  corresponding to the smallest eigenvalue of  $L_1$ Output the phases of the coordinates of  $u$ 

In the experiments that follow, we run the above implementation of polarization using pruning parameters  $\alpha = 0.99$ ,  $\tau = 0.1$ , and the following construction of masks: First, we draw  $K = 3$  masks whose diagonal entries are independent with distribution  $\mathcal{N}(0, 1)$ . Next, we draw  $M - 1$  independent realizations of a random variable  $X$  with uniform distribution over the interval  $[0, 1]$ , and we put  $i \in B$  whenever  $X_i \leq \frac{\log M}{M}$ . This way,  $B \subseteq \mathbb{Z}_M$  is a subset that does not contain 0, but the nonzero members of  $\mathbb{Z}_M$  each reside in  $B$  with probability  $\frac{\log M}{M}$ , thereby following the intent of Theorem 10. Since  $0 \notin B$ , we take  $A = B \cup (-B)$ , and then we construct the auxiliary masks  $\{D_k + \omega^r E^a D_{k'}\}_{r=0}^2$  for every  $(k, k') \in \{0, 1, 2\}$ ,  $k \geq k'$  and  $a \in A$ . Overall, in these experiments, polarization will use a total of  $18|A| + 3$  masks, where  $|A|$  tends to be approximately  $2 \log M$ .

By contrast, PhaseLift offers a lot more flexibility with the number and types of masks used, and this flexibility allows us to run a collection of choice comparisons. We will run experiments at three noise levels, and for each level, we will compare polarization to three different implementations of PhaseLift. In the first implementation, we will only give PhaseLift the measurements corresponding to the original 3 masks we give to polarization. Next, we will give PhaseLift all  $18|A| + 3$  of the masks we described in the previous paragraph. For the last comparison, we will give PhaseLift the original 3 masks along with  $18|A|$  additional masks whose diagonal entries are also independent with distribution  $\mathcal{N}(0, 1)$ . Based on discussions in [7], this last setup appears to be the intended design of Fourier masks for PhaseLift, so in a sense, this last comparison will allow both algorithms to compete with their own “home-field advantage.”

In the comparison between polarization and PhaseLift, we use two performance metrics: run time and relative error of reconstruction, defined in this setting by

$$\min_{\substack{c \in \mathbb{C} \\ |c|=1}} \frac{\|c\hat{x} - x\|_2}{\|x\|_2}.$$

Here,  $c$  is playing the role of the global phase factor we lost in the intensity measurement process. For each noise level tested  $\sigma^2 \in \{0, 0.1, 1\}$ , and for each type of mask ensemble given to PhaseLift (original 3, same  $18|A| + 3$ , random  $18|A| + 3$ ), we considered multiple signal dimensions  $M \in \{2^5, 2^6, 2^7, 2^8, 2^9\}$ . In each of these scenarios, we ran 30 realizations of the following experiment:

- Draw each entry of the signal  $x$  independently from  $\mathcal{N}(0, 1)$ .
- Add independent  $\mathcal{N}(0, \sigma^2)$  noise to each intensity measurement  $|\langle x, \varphi \rangle|^2$ .
- Run both polarization and PhaseLift to estimate  $x$  and record the run time and relative error.

We ran these experiments on an Intel® Core™2 Quad CPU Q9550 at 2.83GHz with 3.6 GB of memory.

Figure 1 presents the results of the noiseless scenarios ( $\sigma^2 = 0$ ). Here, perhaps the most striking distinction between polarization and PhaseLift is the average run time; indeed, polarization produces an estimate in a matter of seconds, whereas PhaseLift typically takes around an hour to terminate in the higher-dimensional scenarios. Letting  $N$  denote the number of intensity measurements, then PhaseLift takes  $\mathcal{O}(N^{3.5})$  operations, whereas the computational bottleneck in polarization is the least-squares step, which takes  $\mathcal{O}(N^3)$  operations, but this does not account for the disparity in run time. Rather, the disparity is indicative of a large constant in front of the  $N^{3.5}$ . In comparing the relative error in reconstruction, we note that PhaseLift terminates when the step size is sufficiently small, so this accounts for the significant difference in performance. Overall, in the noiseless case, polarization handily defeats PhaseLift, even when PhaseLift is given the random masks it prefers.

Next, Figure 2 illustrates how polarization and PhaseLift perform in the presence of some noise ( $\sigma^2 = 0.1$ ). In this case, PhaseLift continues to be rather slow, and it is clear that polarization enjoys greater stability from the  $18|A|$  additional masks when compared to PhaseLift’s performance with just the original 3 random masks. In [7], it is claimed that 3 random masks are sufficient for PhaseLift to reconstruct typical signals (and this is corroborated with our experiments in the noiseless case), but Figure 2 illustrates that PhaseLift requires more masks in order to reconstruct stably. And indeed, when PhaseLift is given more masks, it performs much better. Specifically, when they use the same masks, polarization and PhaseLift produce estimates with statistically comparable relative error, and PhaseLift performs slightly better when given the random masks it prefers. Intuitively, it makes sense that PhaseLift should be at least slightly more stable than polarization since it leverages all of the measurements in a democratic fashion, whereas polarization disenfranchises the vertex and edge measurements that are too small (i.e., “wasting” measurements).

Figure 3 presents the results of our high-noise scenario ( $\sigma^2 = 1$ ), which are very similar to the results corresponding to  $\sigma^2 = 0.1$ .

## 5. DISCUSSION

In this paper, we showed how to leverage the polarization-based phase retrieval technique to construct  $\Theta(\log M)$  Fourier masks that uniquely determine every  $M$ -dimensional signal, and then we used numerical simulations to illustrate the stability of polarization with such masks. At this point, we offer two conclusions:

- (i) The polarization technique for phase retrieval is flexible enough to provably accommodate certain measurement design criteria (such as Fourier masks).
- (ii) If you have the ability to use polarization instead of PhaseLift, you will gain considerable speedups in run time at the price of a slight increase in relative error.

It remains to be rigorously proved that stability holds in the Fourier masks setting. We did not exploit the inherent Fourier structure to further speed up reconstruction, though this could very well be possible. The number of masks might be decreased if additional information about the signal were leveraged, as in [12, 16]. Finally, we still seek Fourier masks-based performance guarantees for PhaseLift and its modifications [11, 21].

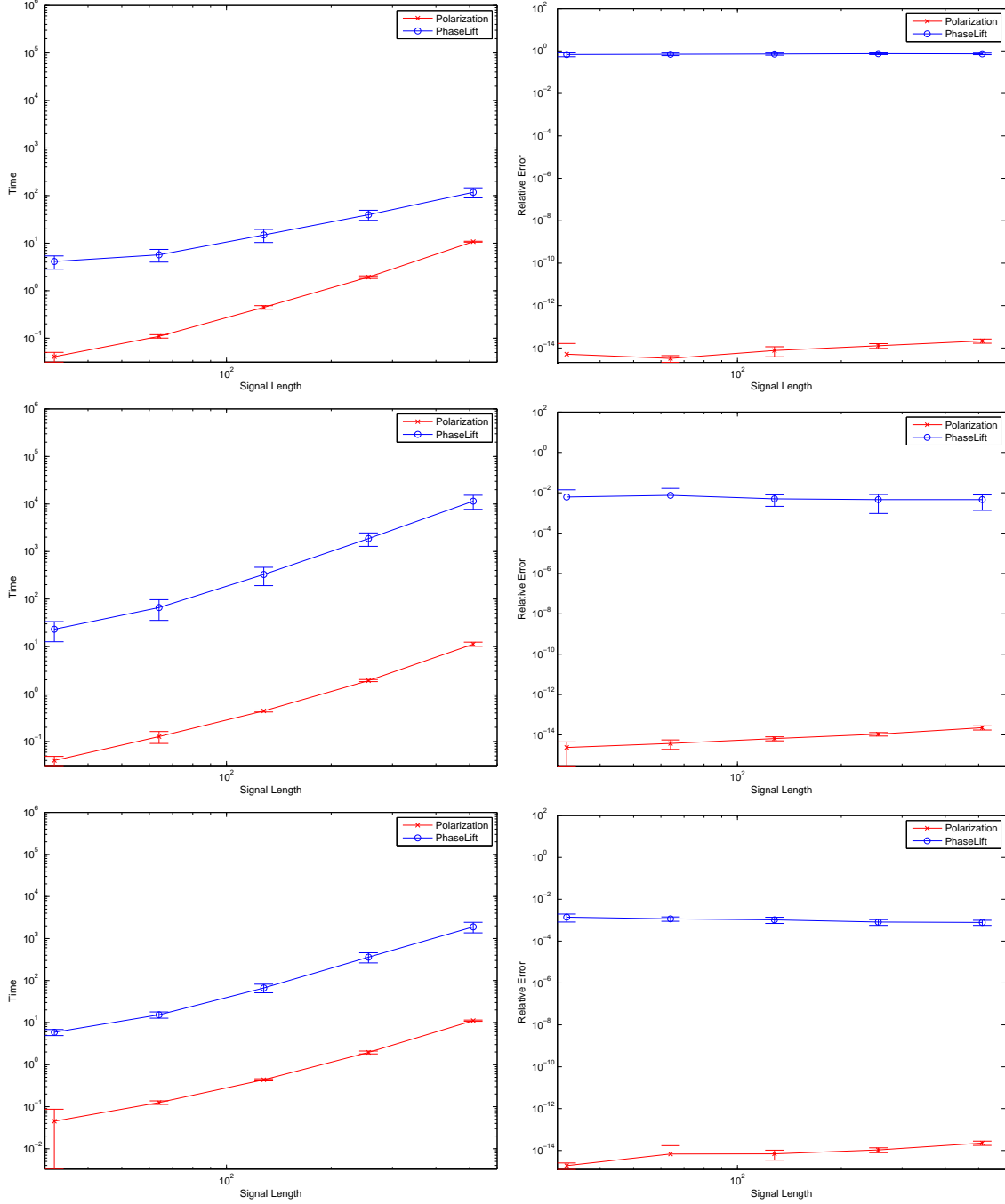


FIGURE 1. Phase retrieval without noise. For the first row of graphs, PhaseLift is given the original  $K = 3$  random masks, for the second, it is given all  $18|A| + 3$  of the masks given to polarization, and in the last row, PhaseLift is given  $18|A| + 3$  random masks. The plotted data illustrate the sample mean plus/minus one sample standard deviation. In some cases, the lower error bar is negative, and so it is not plotted in the log scale.

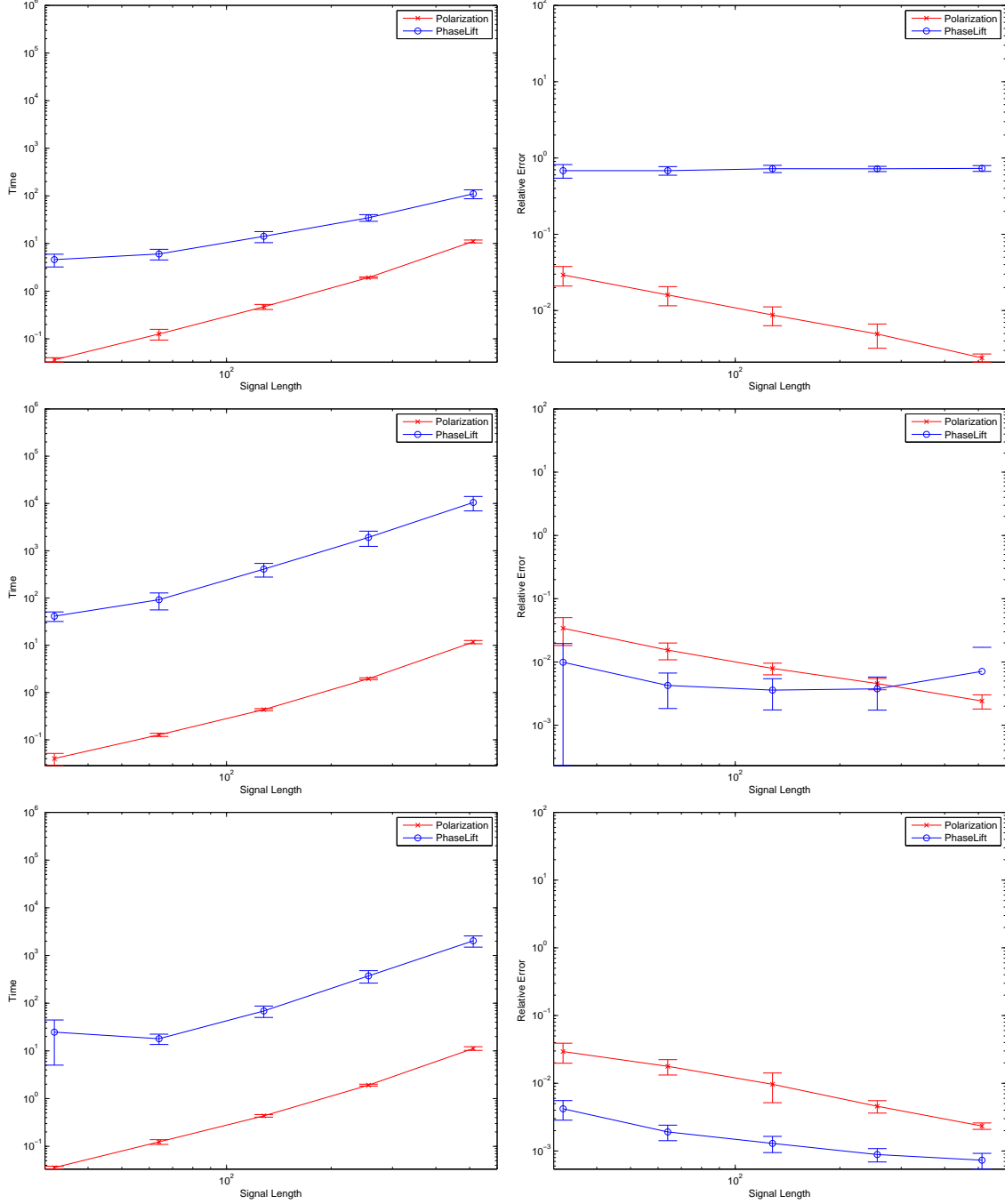


FIGURE 2. Phase retrieval with  $\mathcal{N}(0, 0.1)$  noise added to each intensity measurement. For the first row of graphs, PhaseLift is given the original  $K = 3$  random masks, for the second, it is given all  $18|A| + 3$  of the masks given to polarization, and in the last row, PhaseLift is given  $18|A| + 3$  random masks. The plotted data illustrate the sample mean plus/minus one sample standard deviation. In some cases, the lower error bar is negative, and so it is not plotted in the log scale.

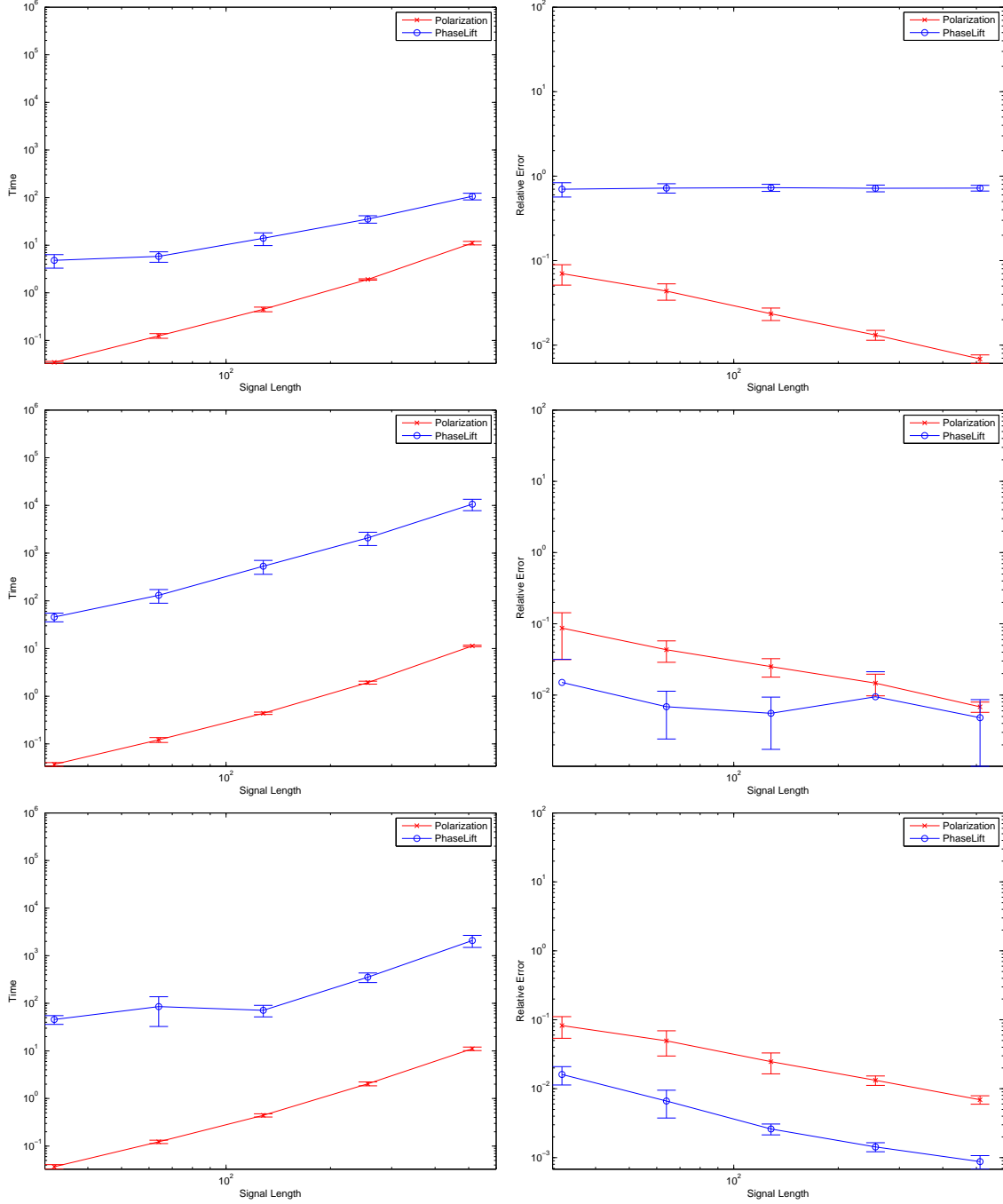


FIGURE 3. Phase retrieval with  $\mathcal{N}(0,1)$  noise added to each intensity measurement. For the first row of graphs, PhaseLift is given the original  $K = 3$  random masks, for the second, it is given all  $18|A| + 3$  of the masks given to polarization, and in the last row, PhaseLift is given  $18|A| + 3$  random masks. The plotted data illustrate the sample mean plus/minus one sample standard deviation. In some cases, the lower error bar is negative, and so it is not plotted in the log scale.



## ACKNOWLEDGMENTS

The authors thank the Erwin Schrödinger International Institute for Mathematical Physics for hosting a workshop on phase retrieval that helped solidify the main ideas in this paper. A. S. Bandeira was supported by NSF DMS-0914892. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

## REFERENCES

- [1] B. Alexeev, A. S. Bandeira, M. Fickus, D. G. Mixon, Phase retrieval with polarization, Available online: arXiv:1210.7752
- [2] B. Alexeev, J. Cahill, D. G. Mixon, Full spark frames, *J. Fourier Anal. Appl.* 18 (2012) 1167–1194.
- [3] N. Alon, Eigenvalues and expanders, *Combinatorica* 6 (1986) 83–96.
- [4] N. Alon, V. D. Milman,  $\lambda_1$ , isoperimetric inequalities for graphs, and superconcentrators, *J. Comb. Theory B* 38 (1985) 73–88.
- [5] A. S. Bandeira, J. Cahill, D. G. Mixon, A. A. Nelson. Saving phase: Injectivity and stability for phase retrieval, Available online: arXiv:1302.4618
- [6] A. S. Bandeira, A. Singer, D. A. Spielman, A Cheeger inequality for the graph connection Laplacian, Available online: arXiv:1204.3873
- [7] E. J. Candès, Y. Eldar, T. Strohmer, V. Voroninski, Phase retrieval via matrix completion, Available online: arXiv:1109.0573
- [8] E. J. Candès, X. Li, Solving quadratic equations via PhaseLift when there are about as many equations as unknowns, Available online: arXiv:1208.6247
- [9] E. J. Candès, T. Strohmer, V. Voroninski, PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming, Available online: arXiv:1109.4499
- [10] J. Cheeger, A lower bound for the smallest eigenvalue of the Laplacian, In: *Problems in analysis (Papers dedicated to Salomon Bochner, 1969)*, Princeton Univ. Press, Princeton, 1970, 195–199.
- [11] L. Demanet, P. Hand, Stable optimizationless recovery from phaseless linear measurements, Available online: arXiv:1208.1803
- [12] A. Fannjiang, Absolute uniqueness of phase retrieval with random illumination, *Inverse Probl.* 28 (2012) 075008.
- [13] R. W. Gerchberg, W. O. Saxton, A practical algorithm for the determination of the phase from image and diffraction plane pictures, *Optik* 35 (1972) 237–246.
- [14] T. Hagerup, C. Rüb, A guided tour of Chernoff bounds, *Inform. Process. Lett.* 33 (1990) 305–308.
- [15] P. Harsha, A. Barth, Lecture 5: Derandomization (Part II), <http://www.tcs.tifr.res.in/~prahladh/teaching/05spring/lectures/lec5.pdf>
- [16] K. Jaganathan, S. Oymak, B. Hassibi, Recovery of sparse 1-D signals from the magnitudes of their Fourier transform, Available online: arXiv:1206.1405
- [17] S. Marchesini, A unified evaluation of iterative projection algorithms for phase retrieval, *Rev. Sci. Instr.* 78 (2007) 011301
- [18] M. Püschel, J. Kovačević, Real, tight frames with maximal robustness to erasures, *Proc. Data Compr. Conf.* (2005) 63–72.
- [19] A. Singer, Angular synchronization by eigenvectors and semidefinite programming, *Appl. Comput. Harmon. Anal.* 30 (2011) 20–36.
- [20] T. Tao, V. H. Vu, *Additive combinatorics*, Cambridge University Press, 2006.
- [21] I. Waldspurger, A. d’Aspremont, S. Mallat Phase recovery, MaxCut and complex semidefinite programming, Available online: arXiv:1206.0102
- [22] J. D. Watson, F. H. C. Crick, A structure for deoxyribose nucleic acid, *Nature* 171 (1953) 737–738.

PROGRAM IN APPLIED AND COMPUTATIONAL MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON,  
NJ 08544

*E-mail address:* `ajsb@math.princeton.edu`

PROGRAM IN APPLIED AND COMPUTATIONAL MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON,  
NJ 08544

*E-mail address:* `yutong@math.princeton.edu`

DEPARTMENT OF MATHEMATICS AND STATISTICS, AIR FORCE INSTITUTE OF TECHNOLOGY, WRIGHT-  
PATTERSON AFB, OH 45433

*E-mail address:* `dustin.mixon@afit.edu`